```
XX
PD   18-APR-1991.
XX
PF   25-SEP-1990;   90WO-US05454.
XX
PR   26-SEP-1989;   89US-0412745.
PR   26-SEP-1990;   90US-0912349.
XX
PA   (TCEL-) T CELL SCI INC.
PA   (UYJO ) JOHNS HOPKINS UNIVERSITY.
PA   (BRIG-) BRIGHAM AND WOMEN'S HOSPITAL.
XX
PI   Fearon DT,  Klickstein LB,  Wong WW,  Carson GR,  Hoh M,  Concino MF;
PI   Makrides SC,  Marsh HC;
XX
DR   WPI; 1991-132854/18.
DR   P-PSDB; AAR11810.
XX
PT   Human complement receptor type 1 gene, encoded proteins and
PT   fragments - for treatment of immune disorders, myocardial infarct,
PT   damage due to inflammmation and in treatment of thrombosis
XX
PS   Claim 1; Fig 1; 234pp; English.
XX
CC   This sequence is a composite of sequences isolated as lambda clones
CC   H10.3, T109.1, H3.1 and H7.1. The clones were present in the
CC   specifically primed lambda gt11 cDNA library (Lambda HH) which was
CC   prepared with cDNA synthesised from poly(A)+ RNA from DMSO induced
CC   HL-60 cells. The library was screened using probes CR1-1, CR-2 and
CC   CR1-4 (see Wong,WW et al.,1985, Proc. Nat. Acad. Sci. USA. 82:7711)
CC   and probe CR1-18 (corresponding to nucleotides 101-352 of this
CC   sequence). There are four direct, long homologous repeats of 450bp,
CC   each comprising 7 short consensus repeats. Nucleotides 28-1533 are
CC   also claimed separately.
CC   See also AAQ11643.
XX
SQ   Sequence 6951 BP; 1802 A; 1681 C; 1659 G; 1809 T; 0 other;


alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                Gaps:      2
 Percent Similarity:  89.865    Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAQ11642   ..

 Align seg 1/1  to: AAQ11642  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGluCysAsnValProGluT 40
         :|||||||||   |||                |||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         |||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::|||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::|||||||||||||   |||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   |||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::|||||||||||||||||:::|||||||:::|||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513


seq_name: /SIDS1/gcgdata/geneseq/genseqn-embl/NA1999.DAT:AAZ38150

seq_documentation_block:
ID   AAZ38150 standard; DNA; 6951 BP.
XX
AC   AAZ38150;
XX
DT   22-FEB-2000  (first entry)
XX
DE   Human C3b/C4b receptor (CR1) protein encoding DNA.
XX
KW   C3B/C4B receptor; CR1 protein; cell-surface protein; erythrocyte; human;
KW   complement regulatory activity; complement pathway enzyme; tissue damage;
KW   reperfusion injury; Arthus reaction; myocardial infarct; inflammation;
KW   heart condition; autoimmune disorder; diagnostic; ss.
XX
OS   Homo sapiens.
XX
PN   US5981481-A.
XX
PD   09-NOV-1999.
XX
PF   06-JUN-1995;   95US-0470652.
XX
PR   03-APR-1989;   89US-0332865.
PR   06-DEC-1974;   74US-0350238.
PR   24-FEB-1993;   93US-0026134.
PR   01-APR-1988;   88US-0176532.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PA   (AVAN-) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI   Concino MF,  Wong WW,  Makrides SC,  Klickstein LB,  Fearon DT,  Ip SH;
PI   Marsh HC,  Carson GR;
XX
DR   WPI; 1999-633357/54.
DR   P-PSDB; AAY55751.
XX
PT   A human C3B/C4B receptor (CR1) protein having antiinflammatory and
PT   cardiant activity -
XX
PS   Disclosure; Fig 1A-P; 87pp; English.
XX
CC   The invention relates to a human C3B/C4B receptor (CR1) protein. The CR1
CC   protein or fragment is expressed as a cell-surface protein on the surface
CC   of a non-human cell and exhibits a complement regulatory activity of full
CC   -length human CR1 as expressed on erythrocytes. The CR1 function in vivo
CC   may be mediated through the inhibition of complement pathway enzymes. The
CC   soluble CR1 protein exhibits a complement regulatory activity, and this
CC   may be used to prevent reperfusion injury, inhibit Arthus reaction, and
CC   neutrophil mediated tissue damage, and reduce myocardial infarct size,
CC   and inflammation. The CR1 protein and its fragments can also be used in
CC   the treatment of conditions which involve unwanted complement activity,
CC   e.g. shock lung, tissue damage due to burn, or ischemic heart conditions,
CC   and autoimmune disorders. CR1 proteins, analogues, derivatives, and anti
CC   -CR1 antibodies are used in assays, and diagnostics. The present sequence
CC   represents a DNA encoding the human CR1 protein.
XX
SQ   Sequence 6951 BP; 1802 A; 1680 C; 1661 G; 1808 T; 0 other;
```

```
alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                Gaps:       2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAZ38150   ..

 Align seg 1/1  to: AAZ38150  from: 1  to: 6951

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||              |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319

      90 ysCysLysArgLysSerCysArgAsnProProAspProValAsnGlyMet 106
         ::|||:::||||||||||||||||||||||||||||||||||||||||||
     320 GGTGCAGACGTAAATCATGTCGTAATCCTCCAGATCCTGTGAATGGCATG 369

     107 AlaHisValIleLysAspIleGlnPheGlySerGlnIleLysTyrSerCy 123
         :::||||||||||||   ||||||||||||||||||||||||||||||||
     370 GTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAAATTAAATATTCTTG 419

     123 sProLysGlyTyrArgLeuIleGlySerSerSerAlaThrCysIleIleS 140
         |   ||||||||||||||||||||||||||||||||||||||||||||||
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::|||||||||||||||||||:::||||||:::||||||
     470 CAGGTGATACTGTCATTTGGGATAATGAAACACCTATTTGTGAC 513
```

Seq_name: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:AAI58380

```
Seq_documentation_block:
ID   AAI58380 standard; cDNA; 7313 BP.
XX
AC   AAI58380;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 583.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM39224.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Claim 1; SEQ ID NO 583; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence 7313 BP; 1903 A; 1770 C; 1733 G; 1907 T; 0 other;
```

```
alignment_scores:
            Quality:  649.50              Length:    148
              Ratio:  4.883                Gaps:       2
 Percent Similarity:  89.865   Percent Identity: 81.757

alignment_block:
 US-10-031-904-8 x AAI58380   ..

 Align seg 1/1  to: AAI58380  from: 1  to: 7313

      10 ProPheProSerArgArgPhe......ProGlyLeuLeuLeuAlaAlaLe 24
         |||   |||:::     |||         |||   |||||||||:::::
      70 CCGGCGCCCGGTCTCCCCTTCTGCTGCGGAGGATCCCTGCTGGCGGTTGT 119

      24 uValLeuLeu...LeuSerSerPheSerAspGlnCysAsnValProGluT 40
         :|||||||||   |||              |||||||||:::|||||||
     120 GGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATGCAATGCCCCAGAAT 169

      40 rpLeuProPheAlaArgProThrAsnLeuThrAspAspPheGluPhePro 56
         ||||||||||||||||||||||||||||||||:::|||||||||||||
     170 GGCTTCCATTTGCCAGGCCTACCAACCTAACTGATGAGTTTGAGTTTCCC 219

      57 IleGlyThrTyrLeuAsnTyrGluCysArgProGlyTyrSerGlyArgPr 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     220 ATTGGGACATATCTGAACTATGAATGCCGCCCTGGTTATTCCGGAAGACC 269

      73 oPheSerIleIleCysLeuLysAsnSerValTrpThrSerAlaLysAspL 90
         ||||||||||||||||||||||||||||||||||:::|||||||||:
     270 GTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGACTGGTGCTAAGGACA 319
```

```
     420 TACTAAAGGATACCGACTCATTGGTTCCTCGTCTGCCACATGCATCATCT 469

     140 erGlyAsnThrValIleTrpAspAsnLysThrProValCysAsp 154
         |||||:::|||||||||||||||||||:::|||||||:::||||||
     470 CAGGTGATACTGTCATTTGCGATAATGAAACACCTATTTGTGAC 513

seq_name: gb_pr:CHPCR1X

seq_documentation_block:
LOCUS       CHPCR1X                 1985 bp    mRNA    linear   PRI 30-OCT-1994
DEFINITION  Pan troglodytes alternatively spliced CR1 (CR1) gene, partial cds.
ACCESSION   L24921
VERSION     L24921.1  GI:557726
KEYWORDS    alternative splicing product; complement receptor 1.
SOURCE      Pan troglodytes cDNA to mRNA.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 1985)
  AUTHORS   Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
            Atkinson,J.P.
  TITLE     Primary sequence of an alternatively spliced form of CR1. Candidate
            for the 75,000 M(r) complement receptor expressed on chimpanzee
            erythrocytes
  JOURNAL   J. Immunol. 153 (2), 691-700 (1994)
  MEDLINE   94292799
FEATURES             Location/Qualifiers
     source          1. .1985
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /cell_line="EBV transformed"
                     /cell_type="B cell"
     sig_peptide     <1. .50
                     /gene="CR1"
     CDS             <1. .>1985
                     /gene="CR1"
                     /codon_start=3
                     /product="complement receptor 1"
                     /protein_id="AAA51439.1"
                     /db_xref="GI:557727"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
                     QIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDRGMTFNLIGESSIRCTSDPQ
                     GNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLYLPGMTISYICDPGYLLVGK
                     GFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMKKVYHYGDYVTLKCEDGYTL
                     EGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIFFILLIIFLSWIILKHRKGN
                     NAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"
     gene            1. .1985
                     /gene="CR1"
     mat_peptide     51. .1985
                     /gene="CR1"
                     /product="complement receptor 1"
BASE COUNT      521 a    476 c    478 g    510 t
ORIGIN


alignment_scores:
            Quality:  645.50            Length:    136
              Ratio:  5.043               Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x CHPCR1X   ..

 Align seg 1/1  to: CHPCR1X  from: 1  to: 1985

      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         ||||||||||::::::||||||||||   |||                |||||

       6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

      35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
         ||||:::|||||||||||||||||||||||||||||||||||||||||:
      56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

      52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
         ::|||||||||||||||||||||||||||||||||||||||||||||||
     106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

      69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
         |||   |||||||||||||||||||||||||||||||||||||||||||
     156 TATTATGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

      85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
         |:::|||||||||:::|||:::||||||||||||||||||||||||||||
     206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

     102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     256 CTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAA 305

     119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
         ||||||||||||||||   ||||||||||||||||||||||||||||||||
     306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

     135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
         |||||||||||||||||:::|||||||||||||||||||:::|||||||:
     356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

     152 alCysAsp 154
         ::||||||
     406 TTTGTGAC 413

seq_name: gb_pr:CHPCR1WT

seq_documentation_block:
LOCUS       CHPCR1WT                6044 bp    mRNA    linear   PRI 30-OCT-1994
DEFINITION  Chimpanzee complement receptor type one (CR1) mRNA, partial cds.
ACCESSION   L24920
VERSION     L24920.1  GI:551564
KEYWORDS    complement receptor 1.
SOURCE      Pan troglodytes cDNA to mRNA.
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1  (bases 1 to 6044)
  AUTHORS   Birmingham,D.J., Shen,X.P., Hourcade,D., Nickells,M.W. and
            Atkinson,J.P.
  TITLE     Primary sequence of an alternatively spliced form of CR1. Candidate
            for the 75,000 M(r) complement receptor expressed on chimpanzee
            erythrocytes
  JOURNAL   J. Immunol. 153 (2), 691-700 (1994)
  MEDLINE   94292799
FEATURES             Location/Qualifiers
     source          1. .6044
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
     gene            1. .6044
                     /gene="CR1"
     CDS             <1. .>6044
                     /gene="CR1"
                     /codon_start=3
                     /product="complement receptor 1"
                     /protein_id="AAA51438.1"
                     /db_xref="GI:557725"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYYGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKDIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPPRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPATPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
```

```
                  QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQ
                  GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYKCRPEYYGRPFS
                  ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHS
                  SAECILSGNAAHWSTKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNPGS
                  RGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFS
                  LNEVVEFRCQPGFAMKGPRRVKCQALNKWEPELPSCSRVCQPPPDVLHAERTQRDKDN
                  FSPGQEVFYSCEPGYDLRGAASLRCTPQGDWSPAAPTCEVKSCDGFMGQLLNGRVLFP
                  VNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCEQIFCPSPPVIPNGR
                  HTGKPLEVFPFGKAVNYTCDPHPDRGTTFDLIGESTIRCTSDPQGNGVWSSPAPRCGI
                  LGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFSITCLDNLVWSSPKD
                  VCRKKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTGHRLIGHSSAECILSGNSAHWS
                  TKPPICQRIPCGLPPTIANGDFISTNRENFHYGSVVTYRCNLGSRGRKVFELVGEPSI
                  YCTSNDDQVGIWSGPAPQCIIPNKCTPPNVENGILVSDNRSLFSLNEVVEFRCQPGFV
                  MKGPHRVKCQALNKWEPELPSCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPG
                  YDLRGAASLHCTPQGDWSPEAPRCAVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCD
                  EGFRLKGSSVSHCVLVGMRSLWNNSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKE
                  ISYTCDPHPDRGMTFNLIGESTIRCTSDPHGNGVWSSPAPRCELPVHAGHCKTPEQFP
                  FASPTIPINDFEFPVGTSLNYECRPGYFGKMFSISCLENLVWSSVEDNCRRKSCGPPP
                  EPFNGMVHINTDTQFGSTVNYSCNEGFRLIGSPSTTCLVSGNNVTWDKKAPICEIISC
                  EPPPTISNGDFYSNNRASFHNGTVVTYQCHTGPDGEQLFELVGERSIYCTSKDDQVGV
                  WSSPPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEIVRFRCQPGFVMVGSHTVQCQT
                  NGRWGPKLPHCSRVCQPPPEILHGEHTPSHQDNFSPGQEVFYSCEPGYDLRGAASLHC
                  TPQGDWSPEAPRCTVKSCDDFLGQLPHGRVLFPLNLQLGAKVSFVCDEGFRLKGRSAS
                  HCVLAGMKALWNSSVPVCEQIFCPNPPAILNGRHTGTPFGDIPYGKEISYACDTHPDR
                  GMTFNLIGESSIRCTSDPQGNGVWSSPAPRCELSVPAACPHPPKIQNGHDIGGHVSLY
                  LPGMTISYICDPGYLLVGKGFIFCTDQGIWSQLDHYCKEVNCSFPLFMNGISKELEMK
                  KVYHYGDYVTLKCEDGYTLEGSPWSQCQADDRWDPPLAKCTSRAHDALIVGTSSGTIF
                  FILLIIFLSWIILKHRKGNNAHEKPKEVAIHLHSQGGSSVHPRTLQTNEENSRVLP"
BASE COUNT     1561 a   1496 c   1462 g   1525 t
ORIGIN


alignment_scores:
            Quality:  645.50            Length:    136
              Ratio:  5.043               Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x CHPCR1WT   ..

 Align seg 1/1  to: CHPCR1WT  from: 1  to: 6044

      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         |||||||||:::::|||||||||    |||                |||||
       6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

      35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
         ||||:::|||||||||||||||||||||||||||||||||||||||||||:
      56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

      52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
         ::|||||||||||||||||||||||||||||||||||||||||||||||||
     106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

      69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
         |||   |||||||||||||||||||||||||||||||||||||||||||||
     156 TATTATGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

      85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
         |:::||||||||:::|||:::|||||||||||||||||||||||||||||
     206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

     102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
         ||||||||||||||:::|||||||||||||||||||||||||||||||||
     256 CTGTGAATGGCATGGTGCATGTGATCAAAGACATCCAATTCGGATCCCAA 305

     119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
         ||||||||||||||||   ||||||||||||||||||||||||||||||||
     306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

     135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
         ||||||||||||||||||:::|||||||||||||||||||:::|||||||:
     356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405
```

```
     152 alCysAsp 154
         ::|||||
     406 TTTGTGAC 413

seq_name: gb_pr:HSCR1RS

seq_documentation_block:
LOCUS       HSCR1RS                2376 bp    mRNA    linear   PRI 22-MAR-1995
DEFINITION  Human CR1 mRNA for C3b/C4b receptor secreted form.
ACCESSION   X14362 Y00812
VERSION     X14362.1  GI:30197
KEYWORDS    alternate splicing; C3b/C4b receptor; complement receptor;
            receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2376)
  AUTHORS   Hourcade,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-NOV-1988) Hourcade D., Howard Hughes Medical
            Institute, 660 S. Euclid St. Louis Mo, 63110, USA
REFERENCE   2  (bases 1 to 2376)
  AUTHORS   Hourcade,D., Miesner,D.R., Atkinson,J.P. and Holers,V.M.
  TITLE     Identification of an alternative polyadenylation site in the human
            C3b/C4b receptor (complement receptor type 1) transcriptional unit
            and prediction of a secreted form of complement receptor type 1
  JOURNAL   J. Exp. Med. 168 (4), 1255-1270 (1988)
  MEDLINE   89010527
COMMENT     The sequence overlaps with that reported by Klickstein et. al. in
            J. Exp. Med. 165:1095-1112(1987) x05309 and in J. Exp. Med.
            168:1699-1717(1988).
FEATURES             Location/Qualifiers
     source          1. .2376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CR1-4"
                     /haplotype="CR1-A"
                     /cell_line="HL-60"
     CDS             <1. .1682
                     /note="CR1 receptor precursor (AA -16 to 543) (1 is 2nd
                     base in codon)"
                     /codon_start=3
                     /protein_id="CAA32541.1"
                     /db_xref="GI:736240"
                     /translation="SLLAVVVLLALPVAWGQCNAPEWLPFARPTNLTDEFEFPIGTYL
                     NYECRPGYSGRPFSIICLKNSVWTGAKDRCRRKSCRNPPDPVNGMVHVIKGIQFGSQI
                     KYSCTKGYRLIGSSSATCIISGDTVIWDNETPICDRIPCGLPPTITNGDFISTNRENF
                     HYGSVVTYRCNPGSGGRKVFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
                     ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGPRRVKCQALNKWEPELPSCSRVCQPPP
                     DVLHAERTQRDKDNFSPGQEVFYSCEPGYDLRGAASMRCTPQGDWSPAAPTCEVKSCD
                     DFMGQLLNGRVLFPVNLQLGAKVDFVCDEGFQLKGSSASYCVLAGMESLWNSSVPVCE
                     QIFCPSPPVIPNGRHTGKPLEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIRCTSDPQ
                     GNGVWSSPAPRCGILGHCQAPDHFLFAKLKTQTNASDFPIGTSLKYECRPEYYGRPFS
                     ITCLDNLVWSSPKDVCKRKSCKTPPDPVNGMVHVITDIQVGSRINYSCTTG"
     sig_peptide     <1. .50
     mat_peptide     51. .1679
                     /product="mature CR1 receptor (AA 1-543); secreted form"
     polyA_site      2376
                     /note="polyA site"
BASE COUNT      633 a    549 c    568 g    626 t
ORIGIN


alignment_scores:
            Quality:  644.50            Length:    136
              Ratio:  5.035               Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x HSCR1RS   ..
```

```
Align seg 1/1  to: HSCR1RS  from: 1  to: 2376

     20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
        ||||||||||::::::||||||||||   |||                ||||| 
      6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

     35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
        ||||:::|||||||||||||||||||||||||||||||||||||||||||:
     56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

     52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
        ::|||||||||||||||||||||||||||||||||||||||||||||||||
    106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

     69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
        ||||||||||||||||||||||||||||||||||||||||||||||||||
    156 TATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

     85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
        |:::||||||||||:::|||:::|||||||||||||||||||||||||||
    206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

    102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
        ||||||||||||||:::|||||||||||||   ||||||||||||||||||
    256 CTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAA 305

    119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
        ||||||||||||||   |||||||||||||||||||||||||||||||||
    306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

    135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
        ||||||||||||||||||:::|||||||||||||||||||:::|||||||:
    356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

    152 alCysAsp 154
        ::|||||||
    406 TTTGTGAC 413

seq_name: gb_sts:G28591

seq_documentation_block:
LOCUS       G28591                2376 bp    DNA     linear   STS 11-JUL-1996
DEFINITION  human STS SHGC-35372, sequence tagged site.
ACCESSION   G28591
VERSION     G28591.1  GI:1408406
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2376)
  AUTHORS   Myers,R.M.
  JOURNAL   Unpublished
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu

            Primer A: TGAGTTGGCAGCAACATCTC
            Primer B: CATACTCTTATATGTGCACTGCCC
            STS size: 216
            PCR Profile:
                    Initial incubation: 94 degrees C for 90 seconds

                    Denaturation:       94 degrees C for 15 seconds
                    Annealing:          62 degrees C for 23 seconds
                    Polymerization:     72 degrees C for 30 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9600

            Protocol:
                    Template:       25 ng
                    Primer:         each 1 uM
                    dNTPs:          each 200 uM
                    Taq Polymerase: 0.05 units/ul
                    Total Vol:      10 ul

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:            50  mM
                    Tris-HCl:       20  mM
                    pH:             8.3

            Prepared with primer pairs provided by Sandoz, derived from X14362
            -- Washington University/Merck EST sequence.
FEATURES             Location/Qualifiers
     source          1. .2376
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="1"
     STS             1680. .1895
     primer_bind     1680. .1699
     primer_bind     complement(1872. .1895)
BASE COUNT      633 a    549 c    568 g    626 t
ORIGIN


alignment_scores:
            Quality:  644.50              Length:    136
              Ratio:  5.036                 Gaps:      1
 Percent Similarity:  94.118   Percent Identity: 86.029

alignment_block:
 US-10-031-904-8 x G28591   ..

 Align seg 1/1  to: G28591  from: 1  to: 2376

      20 LeuLeuAlaAlaLeuValLeuLeu...LeuSerSerPheSerAspGlnCy 35
         ||||||||||::::::||||||||||   |||                ||||| 
       6 CTGCTGGCGGTTGTGGTGCTGCTTGCGCTGCCGGTGGCCTGGGGTCAATG 55

      35 sAsnValProGluTrpLeuProPheAlaArgProThrAsnLeuThrAspA 52
         ||||:::|||||||||||||||||||||||||||||||||||||||||||:
      56 CAATGCCCCAGAATGGCTTCCATTTGCCAGGCCTACCAACCTAACTGATG 105

      52 spPheGluPheProIleGlyThrTyrLeuAsnTyrGluCysArgProGly 68
         ::|||||||||||||||||||||||||||||||||||||||||||||||||
     106 AGTTTGAGTTTCCCATTGGGACATATCTGAACTATGAATGCCGCCCTGGT 155

      69 TyrSerGlyArgProPheSerIleIleCysLeuLysAsnSerValTrpTh 85
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     156 TATTCCGGAAGACCGTTTTCTATCATCTGCCTAAAAAACTCAGTCTGGAC 205

      85 rSerAlaLysAspLysCysLysArgLysSerCysArgAsnProProAspP 102
         |:::||||||||||:::|||:::|||||||||||||||||||||||||||
     206 TGGTGCTAAGGACAGGTGCAGACGTAAATCATGTCGTAATCCTCCAGATC 255

     102 roValAsnGlyMetAlaHisValIleLysAspIleGlnPheGlySerGln 118
         ||||||||||||||:::|||||||||||||   ||||||||||||||||||
     256 CTGTGAATGGCATGGTGCATGTGATCAAAGGCATCCAGTTCGGATCCCAA 305

     119 IleLysTyrSerCysProLysGlyTyrArgLeuIleGlySerSerSerAl 135
         ||||||||||||||   |||||||||||||||||||||||||||||||||
     306 ATTAAATATTCTTGTACTAAAGGATACCGACTCATTGGTTCCTCGTCTGC 355

     135 aThrCysIleIleSerGlyAsnThrValIleTrpAspAsnLysThrProV 152
         ||||||||||||||||||:::|||||||||||||||||||:::|||||||:
     356 CACATGCATCATCTCAGGTGATACTGTCATTTGGGATAATGAAACACCTA 405

     152 alCysAsp 154
         ::|||||||
     406 TTTGTGAC 413
```